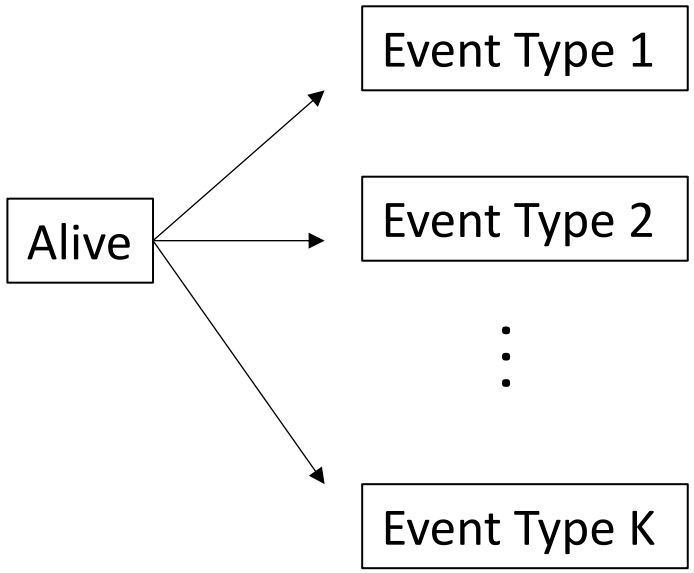
Time-To-Event Analysis in the Presence of Competing risks

Xiaoting (Ting) Wu, Ph.D. MS. , is a biostatistician at the Department of Cardiac Surgery at University of Michigan, US. She has SAS experience for more than 5 years and is one of the awardees for New SAS Professional Award SAS Global Forum 2019. Ting has many year of experience in data and statistical analysis including prediction models, survival analysis, mixed effect models and causal inference. She has broad interests in statistical methodology, statistical consulting, data visualization and SAS applications.

Donald Likosky, Ph.D., is an Associate Professor, and Head of the Section of Health Services Research and Quality in the Depart of Cardiac Surgery at Michigan Medicine. He additionally is the Evaluative Scientist for the MSTCVS Quality Collaborative. He directs the IMPROVE Network (a collaboration of quality collaborative across N. America) and the Perform registry, a perfusion database at over 40 hospitals throughout the United States.

Introduction

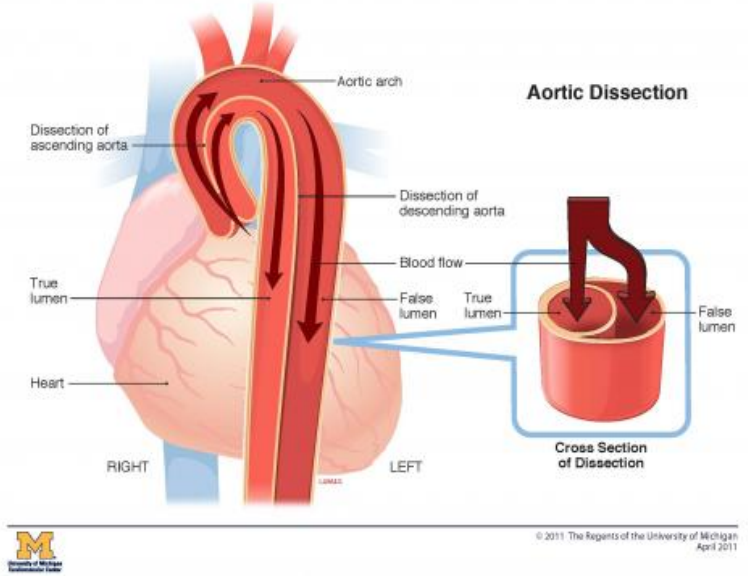
- The presence of competing risk alters the probability of or precludes the occurrence of events of interest. For example, death causes such as stroke, cancer, organ failure are competing events, such that only one of them can occur.

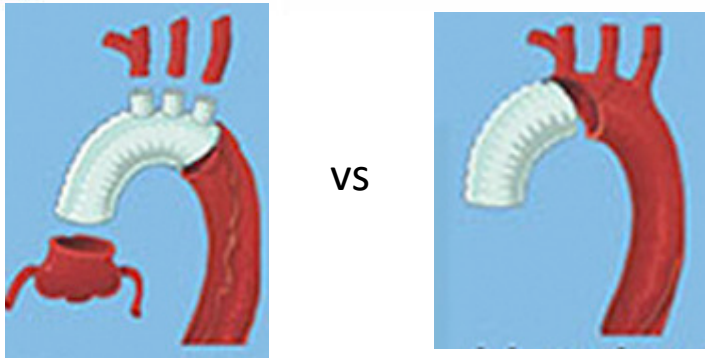


- Competing risks are NOT censoring.

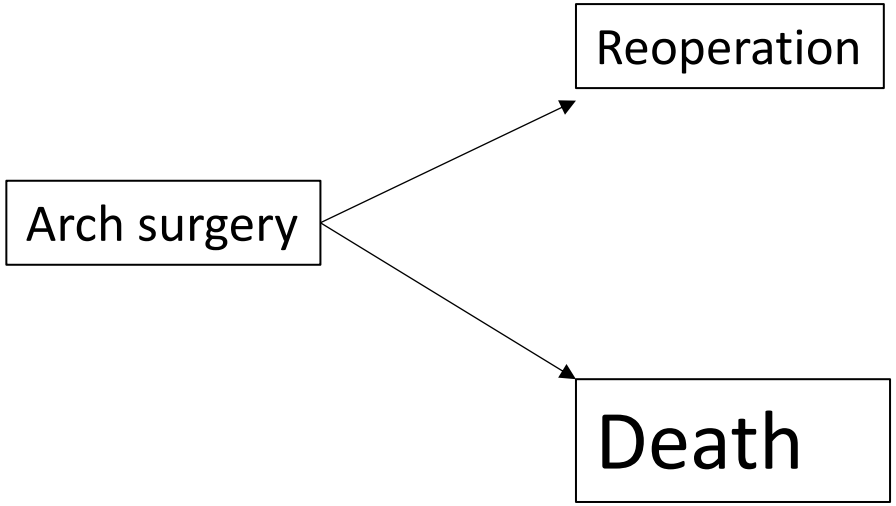
Data example

- The operative outcomes of two different arch replacement surgical strategies are compared – aggressive arch replacement versus conservative hemiarch replacement for patients with acute type A aortic dissection.





vs



Data example

- Event of interest : reoperation
- Competing risk: death
- Time origin: surgery date
- Censoring events: loss of follow up and end of study period.
- Risk factors: age, gender, connective tissue disease status, severe AI condition, and hypertension.

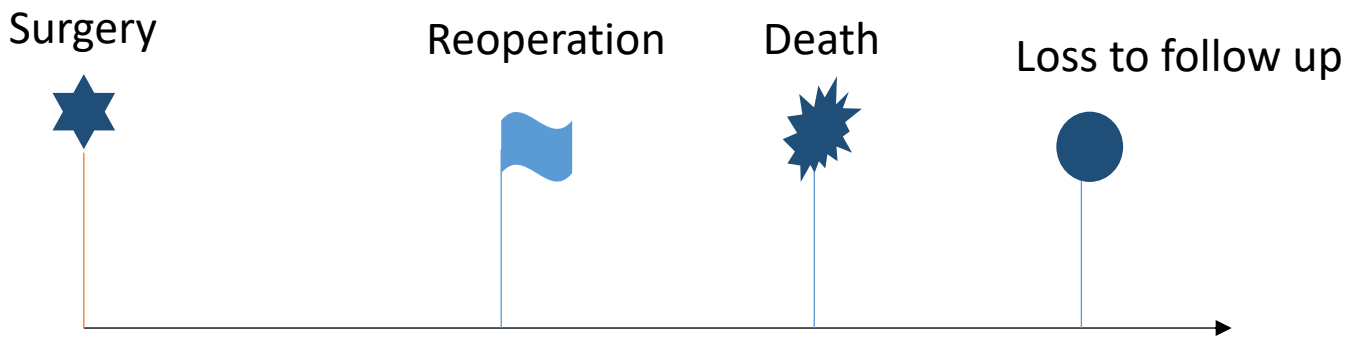
Summary of Failure Outcomes					
Stratum	group	Failed Events	Competing Events	Censored	Total
1	Aggressive arch replacement	15	41	94	150
2	Hemiarch replacement	34	88	200	322
Total		49	129	294	472

Data example

Table 1. Data example variables

Variable name	Variable Meaning
time_reop_arch	Time variable denotes the event time or censor time since surgery
status	0 indicates censor without any event; 1 indicates reoperation; 2 indicates death before arch reoperation
group	1='Aggressive arch replacement' 0='Hemiarch replacement'
sever_AI	Severe aorta insufficiency
age_at_operation	Age at the time of initial operation
gender	Gender 1=female, 0=male
mfs_connect_tissue	Connective tissue disease
htn	Hypertension

Data example



Goal: To analysis reoperation risk over time

Methods in survival analysis with competing risks

- Crude Incidence
- Hazard Function Regression

Crude incidence

- In the absence of competing risks, the cumulative incidence of event can be described as

$$F(t) = 1 - S(t) = \Pr(T \leq t)$$

$S(t)$ is the survival function and can be estimated from Kaplan-Meier $\hat{S}_k(t) = \prod_{t_j \leq t}(1 - \frac{d_{kj}}{n_j})$

- In the presence of competing risks, the cumulative incidence can be described using Cumulative Incidence Function (CIF) with K competing risks. This is interpreted as the probability of experiencing the kth events before time t and before the occurrence of a different type of event

$$F_k(t) = P(T \leq t, \delta = k),\ k = 1, 2, \ldots, K$$

$$\mathrm{CI}(t) = \begin{cases} 0 & \text{if } t \leq t_1 \\ \sum_{t_i \leq t} \left\{ \prod_{j=1}^{i-1} \frac{1 - [d_j + r_j]}{Y_j} \right\} \frac{r_i}{Y_i} & \text{if } t_1 \leq t \end{cases}. \qquad \mathrm{CI}(t) = \sum_{t_i \leq t} \hat{S}(t_i-) \frac{r_i}{Y_i}$$

At time *ti* , let *Yi* be the number of subjects at risk, *ri* be the number of subjects with an occurrence of the event of interest, and *dj* be the number of subjects with an occurrence of competing event. $S(t_i-)$ be the overall survival.

Ref: Lin et al.; Klein 2013

Crude incidence

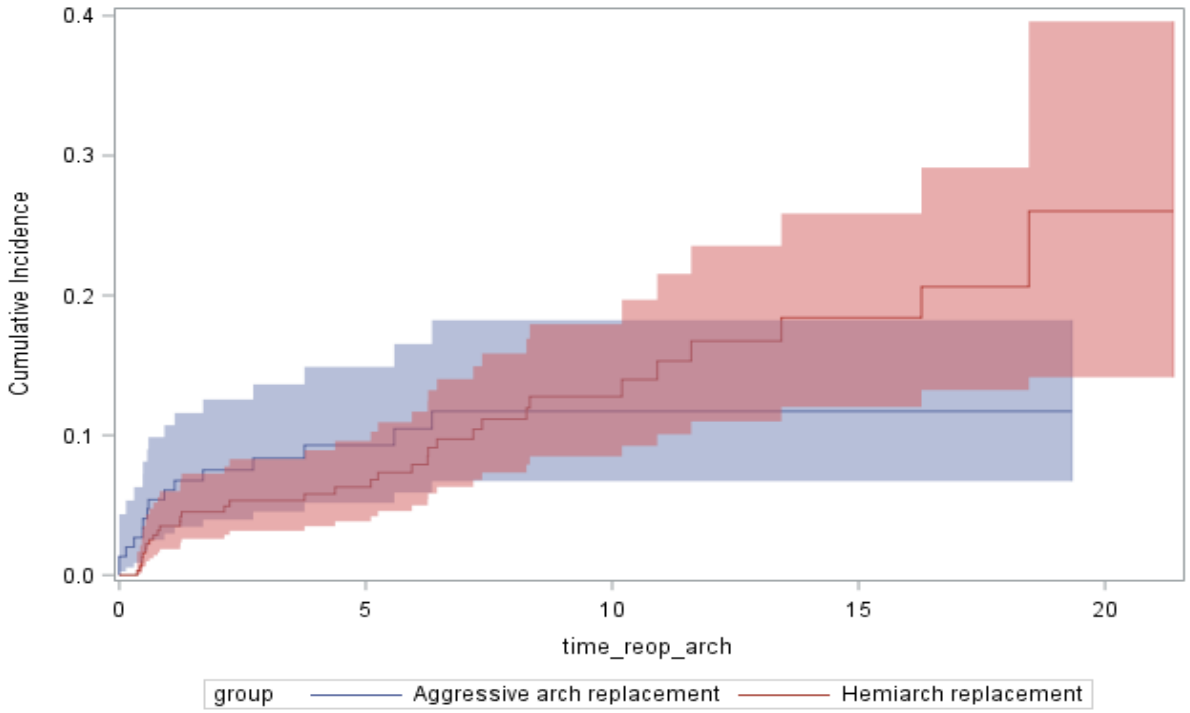
Cautious when estimating incidence function with competing risk:

- When using the complement of Kaplan-Meier, cumulative incidence is greater
- When using the complement of Kaplan-Meier, the sum of the cumulative incidence of each individual outcome will exceed the incidence of the composite outcome of all event types.

Crude incidence

SAS has two equivalent ways to generate cumulative incidence curves: %CIF macro and event codes function in PROC LIFETEST .

```
%CIF (data=arch, time=time_reop_arch, status=status, event=1, censored=0,
group=group, options=plotcl,
 title= CIF macro method);
quit;
```



Crude incidence

- Generate a CIF curve using SAS LIFETEST procedure

```
   * CIF using eventcode option;
 proc lifetest data=arch plots=cif (test cl) atrisk maxtime=18;
       title 'CIF for reoperation risk';
       time time_reop_arch*status(0) /eventcode=1;
       strata group;
 run;
```

- Compare to KM estimates using SAS LIFETEST procedure

```
proc lifetest  data=arch outsurv=km_sur2 plots=survival(cl test);
       time time_reop_arch*status (0,2) ;
       strata group;
run;
```

Crude incidence

From SAS output, we could obtain the cumulative incidence over time. Here is an example output.

Stratum 1: group = 0				
time_reop_arch	Cumulative Incidence	Standard Error	95% Confidence Interval	
0	0	0	.	.
0.364384	0.00318	0.00318	0.000308	0.0167
0.419178	0.00636	0.00449	0.00129	0.0213
0.452055	0.00954	0.00549	0.00266	0.026
0.465753	0.0127	0.00633	0.00427	0.0305
0.471233	0.0159	0.00707	0.00604	0.0349
0.531507	0.0191	0.00774	0.00793	0.0392
……				

Comparisons of CIF from Kaplan-Meier method versus Subdistribution method

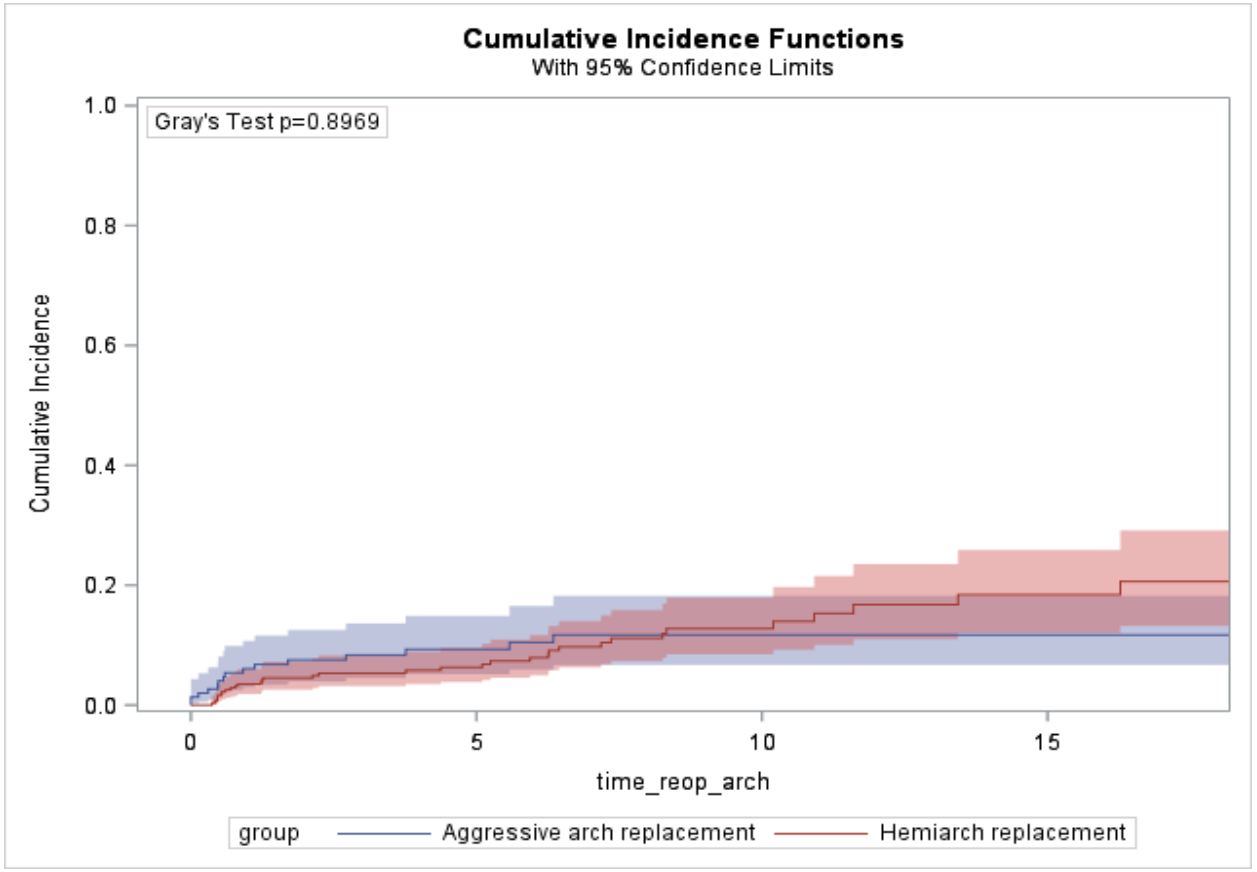
Table 2. Comparisons of cumulative incidence from the complement of Kaplan-Meier method versus CIF

Group	Types of events	5-year incidence from (1-S(t))	5-year CIF
0	Composite (death, reop)	0.250	0.250
1	Composite (death, reop)	0.318	0.318
0	death	0.193	0.1868
1	death	0.238	0.2251
0	reop	0.071	0.063
1	reop	0.105	0.093
0	Sum (death, reop)	0.264	0.250
1	Sum (death, reop)	0.343	0.318

- When using the complement of Kaplan-Meier, cumulative incidence is greater.
- When using the complement of Kaplan-Meier, the sum of the cumulative incidence of each individual outcome will exceed the incidence of the composite outcome of all event types.

Crude incidence

Gray's test (`plots=cif (test cl))` could be used to test the difference of cumulative incidence curve in the two groups. In the absence of censoring, Gray's test (Gray (1988)) is identical to the log-rank test. The two tests differ in the presence of competing risk.



Hazard function regression

- In the absence of competing risks, the hazard function describes the instantaneous rate of occurrence of the event of interest in subjects who're still at risk of event

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t}.$$

- COX proportional hazard regression model

$$\lambda(t) = \lambda_0(t)\exp(\mathbf{X}\boldsymbol{\beta}) \quad \text{Or} \quad \log(\lambda(t)) = \log(\lambda_0(t)) + \mathbf{X}\boldsymbol{\beta}$$

$\lambda_0(t)$ defines the baseline hazard function, X is the set of variables, and β is the regression parameter.

Hazard ratio is the exponential of the regression coefficient and can be interpreted as the relative change in hazard associated with a unit change in the predicator variables.

Ref: Austin, Circulation 2016

Hazard function regression

- In the presence of competing risks, the hazard function can be expressed as cause specific hazard function and subdistribution hazard function.
- The cause-specific hazard function can be interpreted as the instantaneous rate of occurrence of the kth event In subjects who have not yet experienced any of the different types of events.

$$\lambda_k^{cs}(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}(t \le T < t + \Delta t, D = k \,|\, T \ge t)}{\Delta t}.$$

- Modeling the cause specific hazard: $h_1(t|\mathbf{Z}) = h_{10}(t)\exp(\beta'\mathbf{Z})$

The cause-specific cox regression is recommended for studying risk factor effect. The parameters are estimated by maximizing partial likelihood. The risk set exclude those who have previously experienced a competing event during model estimates.

$$L(\beta) = \prod_i \left(\frac{\exp(\beta'\mathbf{Z}_i)}{\sum_{j \in \mathcal{R}_i} \exp(\beta'\mathbf{Z}_j)} \right)^{\delta_i = 1}$$

$\mathcal{R}_i$ is the risk set of patients who do not fail or are not censored before Xi .

Hazard function regression

- The subdistribution hazard function (Gray's method) can be interpreted as the instantaneous risk of occurrence of the kth event in subjects who have not yet experienced kth types of events.

$$\lambda_k^{sd}(t)=\lim_{\Delta t\to 0}\frac{\text{Prob}(t<T\le t+\Delta t, D=k \mid T>t\cup(T<t\cap K\neq k))}{\Delta t}$$

- Modeling the cumulative incidence (Fine and Gray (1999)) $\tilde{h}_1(t|\mathbf{Z})=\tilde{h}_{10}(t)\exp(\beta'\mathbf{Z})$

The CIF regression model is recommended for risk prediction. The parameters are estimated by maximizing partial likelihood. The risk set includes those who have previously experienced a competing event.

$$\tilde{L}(\beta)=\prod_i\left(\frac{\exp(\beta'\mathbf{Z}_i)}{\sum_{j\in\tilde{\mathcal{R}}_i} w_{ij}\exp(\beta'\mathbf{Z}_j)}\right)^{\delta_i=1}$$

$\tilde{\mathcal{R}}_i$ includes patients at risk for event of interest and patients with a competing event before time Xi. Weights $w_{ij}=1$ is given for patients with no event of interest before time Xi; while weight w_{ij} that reduces with time is given for patients with competing risk.

Cause specific hazard model

To fit a cause specific hazard model, the competing risk is treated as a censoring event, so status (0,2) indicated that both alive without reoperation, and death before any reoperation are treated as censoring in the model. Treating all competing events as censoring ensures that the risk set at each event time contains only those subjects who did not experience any competing events or are truly censored. The existing tools such as ASSESS statement can be used to check model assumptions. Starting in SAS/STAT 14.3, you may also use EVENTCODE (COX)=option in the MODEL statement to fit the cause-specific Cox models.

```
* cause-specific using PHREG;
proc phreg data=arch;
class group (ref="0") gender  sever_AI(ref="0") mfs_connect_tissue (ref="0") htn
(ref="0");
model time_reop_arch*status(0,2)=group age_at_operation gender  sever_AI
mfs_connect_tissue htn;
hazardratio group/diff=ref;
hazardratio age_at_operation/units=10;
hazardratio gender/diff=ref;
hazardratio sever_AI/diff=ref;
hazardratio mfs_connect_tissue/diff=ref;
hazardratio htn/diff=ref;
run;
```

Subdistribution hazard model

To fit a subdistribution model, we could use eventcode option in the model statement in PHREG procedure. Here, event code=1 indicated that reoperation is the event of interest, 0 is alive without reoperation, and coding 2 is the competing risk of death. For this Fine and Gray model, you could predict CIFs for the event using BASLINE statement.

```
* subdistribution using PHREG;
proc phreg data=arch plots(overlay=bystratum)=cif ;
class group (ref="0") gender  sever_AI(ref="0") mfs_connect_tissue
(ref="0") htn (ref="0");
model time_reop_arch*status(0)=group age_at_operation gender  sever_AI
mfs_connect_tissue htn/eventcode=1;
hazardratio group/diff=ref;
hazardratio age_at_operation/units=10;
hazardratio gender/diff=ref;
hazardratio sever_AI/diff=ref;
hazardratio mfs_connect_tissue/diff=ref;
hazardratio htn/diff=ref;
run;
```

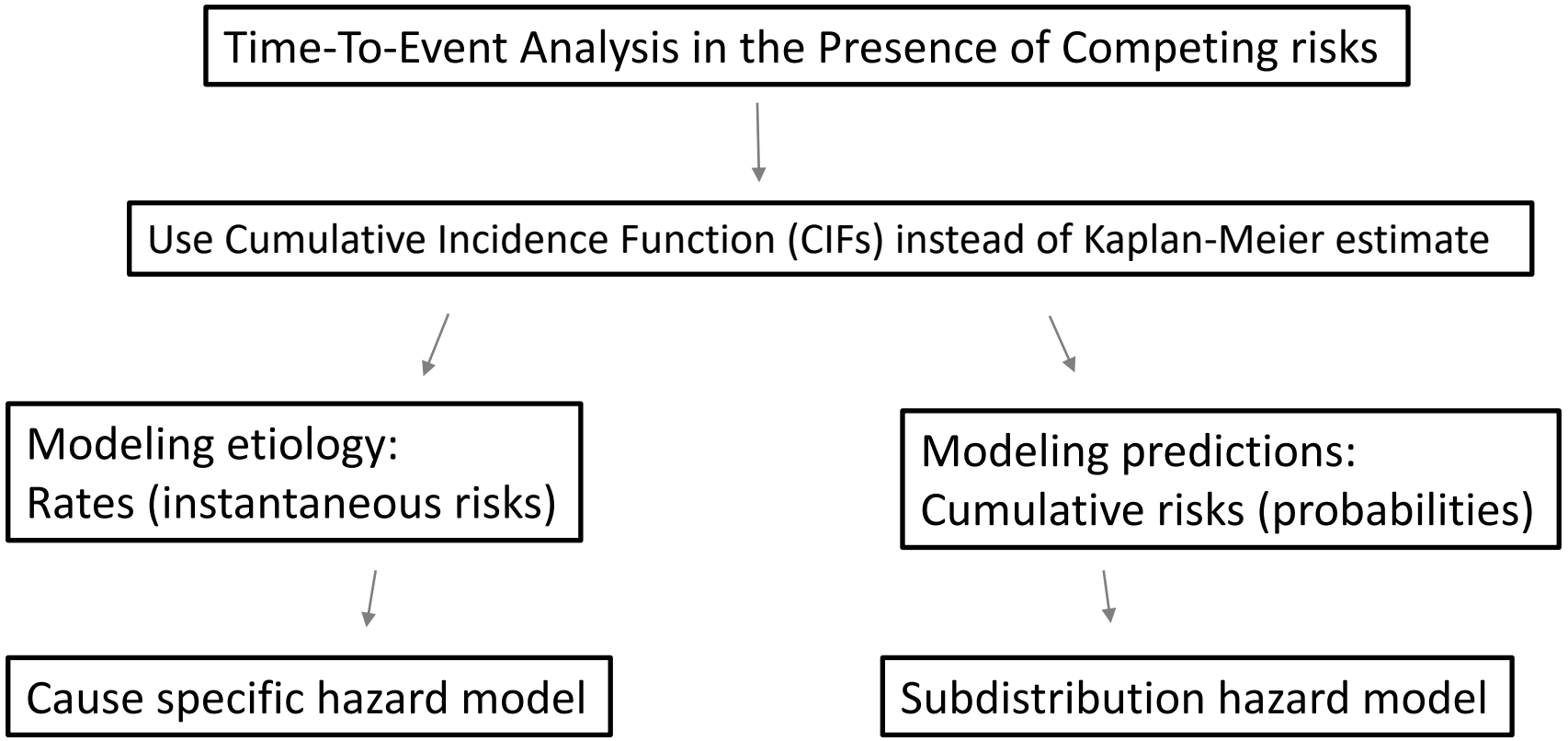
Comparisons of two hazard models

Methods	Subdistribution			Cause-Specific			Regular COX		
Event of interest	Reoperation			Reoperation			Death		
Risk factors	Hazard ratio	95% Wald Confidence Limits		Hazard ratio	95% Wald Confidence Limits		Hazard ratio	95% Wald Confidence Limits	
Group 1 vs 0	0.88	0.47	1.62	0.87	0.47	1.62	1.16	0.80	1.68
Age at operation Unit=1	0.97	0.96	1.00	0.98	0.96	1.01	1.04	1.03	1.06
Age at operation Unit=10	0.77	0.63	0.95	0.84	0.66	1.06	1.51	1.29	1.76
Gender 1 vs 2	1.46	0.72	2.99	1.49	0.73	3.04	1.06	0.72	1.56
Sever_AI 1 vs 0	0.46	0.23	0.91	0.43	0.19	0.95	1.05	0.70	1.59
Connect tissue disease 1 vs 0	1.13	0.51	2.49	1.22	0.43	3.48	1.52	0.63	3.67
Hypertension 1 vs 0	1.10	0.60	2.02	1.09	0.59	2.03	1.03	0.69	1.54

Interpretation example: Age is a pronounced risk factor for death. A 10-year increase in age increases the hazard of death by 51% (HR=1.51, 95% CI (1.29, 1.76)). This impacts on the occurrence of reoperation. A 10-year increase in age decreased the relative incidence of reoperation by 23% (HR=0.77, 95% CI (0.63, 0.95)), and it decreased cause-specific hazard of reoperation by 16% (HR=0.84, 95%CI (0.66, 1.06)).

Conclusion

This paper demonstrates SAS applications for cumulative incidence function and cause-specific hazard function in time-to-event analysis adjusting for competing risk events.



Reference

1. Austin, P.C., D.S. Lee, and J.P. Fine, *Introduction to the Analysis of Survival Data in the Presence of Competing Risks.* Circulation, 2016. **133**(6): p. 601-9.
2. John P. Klein , M.L.M., *Survival analysis techniques for censored and truncated data*. 2003.
3. Fine, J.P., and Gray, R. J., *A Proportional Hazards Model for the Subdistribution of a Competing Risk.* Journal of the American Statistical Association, 1999. **94**: p. 496–509.
4. Guixian Lin, Y.S., Gordon Johnston, *Analyzing Survival Data with Competing Risks Using SAS® Software.* SAS Global Forum 2012, 2012.
5. Ying So, G.L., and Gordon Johnston, *Using the PHREG Procedure to Analyze Competing-Risks Data.* SAS Institute Inc., 2014.
6. So, C.G.a.Y., *Cause-Specific Analysis of Competing Risks Using the PHREG Procedure.* SAS Institute Inc., 2018.

Thank you!

Xiaoting Wu (Ting), PhD, MS
Department of Cardiac Surgery
University of Michigan
Work Phone: 734.936.7731
E-Mail: xiaotinw@med.umich.edu